by either theory. Without going into detail, we have found a perfect match in that region with respect to the amplitude of the weight fluctuations, particularly its decrease along the tapped-delay line in the third-order approximation.

IV. Conclusions

This correspondence deals with the correlation matrix V of the weight errors in an LMS-type adaptive TDL filter. Restricting ourselves to a white output noise and avoiding any independence assumption, we have determined the coefficients of a power series $V = V_1\mu + V_2\mu^2 + V_3\mu^3 + \cdots$ in terms of the stepsize μ. The first term $V_1\mu$ is a scalar matrix, representing a set of equal-power, uncorrelated weight fluctuations, in agreement with what is found with the aid of the independence assumption [1]. The quadratic approximation $V_1\mu + V_2\mu^2$ represents a set of weakly correlated equal-power weight fluctuations with a slightly increased common power level. In the third-order approximation, we observe a power decrease along the delay line. This effect can run up to several percent and is more easily observed than the second-order effects [8].

We expect that the proposed iterative method will also lend itself to the treatment of adjacent questions such as adaptation transients and filter tracking. In addition, it might be applicable to other adaptive algorithms like the normalized LMS type. We were able to show that an independence assumption is not required so that teaching adaptive filtering is released from an inconsistent tool [11].

Acknowledgment

Discussions with W. F. G. Mecklenbräuker, G. Kubin, and S. Haykin are gratefully acknowledged. Their critical remarks have led to a substantial improvement of the manuscript.

References

[1] S. Haykin, *Adaptive Filter Theory*, 3rd ed. London, U.K.: Prentice-Hall, 1996.

[2] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[3] V. Solo, "The error variance of LMS with time-varying weights," *IEEE Trans. Signal Processing*, vol. 40, pp. 803–813, 1992.

[4] S. C. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. Signal Processing*, vol. 43, pp. 2863–2871, 1995.

[5] H. J. Butterweck, "A steady-state analysis of the LMS adaptive algorithm without use of the independence assumption," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Detroit, MI, May 9–12, 1995, pp. 1404–1407.

[6] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes With Applications to Stochastic System Theory*. Cambridge, MA: MIT Press, 1984.

[7] H. J. Butterweck, "An approach to LMS adaptive filtering without use of the independence assumption," in *Proc. Eur. Conf. Signal Processing*, G. Ramponi, Ed., Trieste, Italy, 1996, pp. 1223–1226.

[8] ——, "Iterative analysis of the steady-state weight fluctuations in LMS-type adaptive filters," Tech. Rep. 96-E-299, Eindhoven Univ. Technol., Eindhoven, The Netherlands, June 1996.

[9] S. Florian and A. Feuer, "Performance analysis of the LMS algorithm with a tapped delay line (two-dimensional case)," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1542–1549, 1995.

[10] S. C. Douglas and T. H.-Y. Meng, "Exact expectation analysis of the LMS adaptive filter without the independence assumption," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 1992, San Francisco, CA, vol. IV, pp. 61–64.

[11] H. J. Butterweck, "The independence assumption: A dispensable tool in adaptive filter theory," *Signal Process.*, vol. 57, pp. 305–310, 1997.

Sufficient Stability Bounds for Slowly Varying Direct-Form Recursive Linear Filters and Their Applications in Adaptive IIR Filters

Alberto Carini, V. John Mathews, and Giovanni L. Sicuranza

***Abstract*—This correspondence derives a sufficient time-varying bound on the maximum variation of the coefficients of an exponentially stable time-varying direct-form homogeneous linear recursive filter. The stability bound is less conservative than all previously derived bounds for time-varying IIR systems. The bound is then applied to control the step size of output-error adaptive IIR filters to achieve bounded-input bounded-output (BIBO) stability of the adaptive filter. Experimental results that demonstrate the good stability characteristics of the resulting algorithms are included. This correspondence also contains comparisons with other competing output-error adaptive IIR filters. The results indicate that the stabilized method possesses better convergence behavior than other competing techniques.**

***Index Terms*—Adaptive IIR filter, time-varying recursive linear filter.**

I. Introduction

Adaptive IIR filters have been the subject of active research over the last three decades [5], [9], [11], [12], [15]. Despite a large amount of work that has been done, some open issues still remain. One of these issues is that of ensuring the stability of the time-varying IIR filter that results from the identification process.

Researchers have attempted to derive adaptive IIR filters that operate in a stable manner in several different ways. One class of algorithms is obtained by means of the equation-error technique. In the equation-error technique, the IIR filter is identified by the use of a two-channel adaptive FIR filter that operates on samples of the input and the desired response signals. Since the system model employed in equation-error methods is not recursive, the adaptive filter can operate in a stable manner when the step size is properly selected. However, this fact does not ensure the stability of the resulting IIR filter. Moreover, it is well-known that equation-error adaptive algorithms give biased solutions when the desired response signal is corrupted by noise.

Output error algorithms have become popular in adaptive IIR filtering research in recent years. In output error techniques, the adaptive filter operates in a recursive manner on the input signal to provide an estimate of the desired response signal. A class of such methods requires a certain system transfer function to be strictly positive real (SPR) in order to avoid problems with instability and to ensure the convergence of the algorithm. This class of algorithms includes the pseudo-linear regression algorithm (PRA) [3], which is also known as Feintuch's algorithm, Landau's algorithm [7], the hyperstable adaptive recursive filter (HARF) [4], and the simplified

Manuscript received January 29, 1997; revised February 16, 1999. This work was supported in part by NATO under Grant CRG.950379 and Esprit LTR Project 20229 Noblesse. Parts of this paper were presented at ICASSP, Munich, Germany, April 1997. The associate editor coordinating the review of this paper and approving it for publication was Dr. Mahmood R. Azimi-Sadjadi.

V. J. Mathews is with the Department of Electrical Engineering, University of Utah, Salt Lake City, UT 84112 USA.

A. Carini is with Telital S.p.A., Trieste, Italy (e-mail: alberto.carini@telital.it).

G. L. Sicuranza is with the DEEI, Università degli Studi di Trieste, Trieste, Italy (e-mail: sicuranza@univ.trieste.it).

Publisher Item Identifier S 1053-587X(99)06749-5.

HARF (SHARF) [8]. An SPR condition is not easy to guarantee in practice. In the PRA, the SPR condition limits the range of the location of the poles of the unknown system for which convergence is guaranteed. This problem is avoided in HARF and SHARF algorithms. However, some *a priori* knowledge of the underlying system model is required in order to meet the SPR condition. Moreover, if the system consists of two or more parallel sections, the SPR condition is not sufficient to guarantee stability [16]. A second class of adaptive output-error direct-form filters employ stability monitoring by checking the location of the instantaneous poles of the system and projecting the coefficients back to a region for which the instantaneous poles are within the unit circle [9]. Unfortunately, time-varying filters may be unstable even when the instantaneous poles are within the unit circle. A simple example is given by the time-varying recursive linear system with two coincident poles located at $(-1)^{k-1}0.5$ at time k with input–output relationship

$$y(k) = (-1)^{k-1}y(k-1) - 0.25y(k-2) + x(k). \tag{1}$$

Even though the instantaneous poles are always bounded by one and far from the unit circle itself, it is straightforward to show that the response of this system to a unit impulse signal diverges exponentially. Consequently, even though projection-based techniques that force the instantaneous poles of the system to stay within the unit circle work well in a large number of situations, they are not guaranteed to operate in a stable manner in all situations. A third class of output-error algorithms employs lattice structures [10], [12]. Normalized lattice filters are guaranteed to be stable if the reflection coefficients are bounded by one. Similar conditions can also be established for other filter structures such as power wave digital filters [6] and normal forms [12]. However, direct-form filters are particularly suited for the multiply-accumulate architectures found in most digital signal processors and, for this reason, are often preferred to the above-mentioned filter structures.

In this correspondence, we present a method for controlling the adaptation step size to guarantee bounded-input bounded-output stability of output-error adaptive IIR filters. It is well-known [1], [2], [12]–[14] that a recursive time-varying homogeneous linear system is exponentially stable if its instantaneous poles are always inside the unit circle and if they are sufficiently slowly varying. We first derive a new upper bound on the maximum allowable coefficient variation for the stability of a direct-form linear recursive filter and then apply the results to control the step size of an adaptive IIR filter to ensure stable operation. Experimental results demonstrating the good convergence characteristics of the adaptive filter so derived, as well as comparing our stability bound with previously available results, are also included.

II. Sufficient Stability Conditions for Slowly-Varying Direct-Form Recursive Systems

We consider a time-varying recursive linear system with input–output relationship given by

$$y(k) = \sum_{i=0}^{N-1} b_i(k)x(k-i) + \sum_{i=1}^{N-1} a_i(k)y(k-i). \tag{2}$$

Let

$$\begin{aligned}\boldsymbol{\theta}(k) = [&b_0(k), b_1(k), \cdots, b_{N-1}(k), a_1(k)\\ &a_2(k), \cdots, a_{N-1}(k)]^T\end{aligned} \tag{3}$$

denote the coefficient vector, and let the evolution of the coefficients be of the form

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \mu_k \boldsymbol{\psi}(k) \tag{4}$$

where μ_k is a time-varying scalar sequence. Our objective is to find a sufficient bound on the squared-norm of the increment vector $\mu_k\boldsymbol{\psi}(k)$ given by $\mu_k^2\psi^T(k)\psi(k)$ such that the time-varying system of (2) is stable in the bounded-input, bounded-output (BIBO) sense. From such a result, we then find a bound on μ_k for guaranteeing the stability of the system. An adaptive filter with coefficient update as in (4) will be BIBO stable if μ_k is chosen smaller than or equal to such a bound. The basis for our work is the following theorem proved in [14]:

Theorem 1: The linear state equation

$$\boldsymbol{x}(k+1) = \boldsymbol{A}(k)\boldsymbol{x}(k), \qquad \boldsymbol{x}(k_0) = \boldsymbol{x}_0 \tag{5}$$

is uniformly exponentially stable if and only if there exists an $N \times N$ matrix sequence $\boldsymbol{Q}(k)$ that is symmetric for all k and such that

$$\eta\boldsymbol{I} \le \boldsymbol{Q}(k) \le \rho\boldsymbol{I} \tag{6}$$

and

$$\boldsymbol{A}^T(k+1)\boldsymbol{Q}(k+1)\boldsymbol{A}(k+1) - \boldsymbol{Q}(k) \le -\gamma\boldsymbol{I} \tag{7}$$

where η, ρ, and γ are finite positive constants.

The condition "matrix $\boldsymbol{Q} \le \rho\boldsymbol{I}$" in the theorem implies that $\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \le \rho\boldsymbol{x}^T\boldsymbol{x}$ for all vectors $\boldsymbol{x}$. Exponential stability of the homogeneous system implies BIBO stability of the more general system in (2), provided that the coefficients of the nonrecursive part are bounded [12].

Theorem 1 is expressed in terms of the state transition matrix $\boldsymbol{A}(k)$ while we are interested in the direct-form realization. However, it is trivial to transform the direct-form representation in (2) to the state space representation by considering the state vector

$$\boldsymbol{x}(k) = [y(k-1), y(k-2), \cdots, y(k-N)]^T \tag{8}$$

and the corresponding state transition matrix

$$\boldsymbol{A}(k) = \begin{bmatrix} a_1(k) & a_2(k) & \cdots & a_{N-1}(k) & a_N(k) \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{9}$$

Finite and bounded choices of the feedforward coefficients $b_i(k)$ do not affect the stability of the system. Consequently, ignoring these coefficients in the rest of the analysis does not result in any loss of generality. In the derivations that follow, we find a sequence of candidate matrices $\boldsymbol{Q}(k)$ that meets the Lyapunov conditions given by (6) and (7) for slowly varying recursive linear filters. We assume that the instantaneous poles of the recursive system are always inside the unit circle.

A. Lyapunov Candidate Sequence

Since the poles of the system (2) are by hypothesis always inside the unit circle, we can consider as the Lyapunov candidate the unique, symmetric, and positive definite solution of the discrete-time Lyapunov equation

$$\boldsymbol{A}^T(k)\boldsymbol{Q}_{\nu_k}(k)\boldsymbol{A}(k) - \boldsymbol{Q}_{\nu_k}(k) = -\nu_k\boldsymbol{I}_N \tag{10}$$

where ν_k is a bounded positive sequence with $\nu_k \in [\nu_{\min}, \nu_{\max}]$, and $[\nu_{\min}, \nu_{\max}]$ is an arbitrary interval of the positive real axis.

Following the derivations in [2] and [14], it can be shown that $\boldsymbol{Q}_{\nu_k}(k)$ is given by

$$\boldsymbol{Q}_{\nu_k}(k) = \nu_k \sum_{i=0}^{+\infty} [\boldsymbol{A}^T(k)]^i \boldsymbol{A}^i(k). \tag{11}$$

The closed-form solution for $\boldsymbol{Q}_{\nu_k}(k)$ is given by

$$\mathrm{vec}[\boldsymbol{Q}_{\nu_k}(k)] = -\nu_k[\boldsymbol{A}^T(k) \otimes \boldsymbol{A}^T(k) - \boldsymbol{I}_{N^2}]^{-1}\ \mathrm{vec}[\boldsymbol{I}_N] \tag{12}$$

where $\otimes$ indicates the Kronecker product, and "vec[$\boldsymbol{Q}$]" is a vector operator that stores the columns of $\boldsymbol{Q}$ in a predetermined order.

Since ν_k is bounded by the positive and finite values $\nu_{\min}$ and $\nu_{\max}$ and the eigenvalues of $\boldsymbol{A}(k)$ are in the open unit disk of the complex plane, it can be shown that the Lyapunov conditions in (6) are always satisfied. This result can be proved easily by following the derivations in [2] and [14]. As for the condition in (7), let us consider (10) at time $k+1$ and add $\boldsymbol{Q}_{\nu_{k+1}}(k+1) - \boldsymbol{Q}_{\nu_k}(k)$ to both sides of the expression. It follows trivially that the condition of (7) can be rewritten as

$$\begin{aligned} \boldsymbol{A}^T(k+1)\boldsymbol{Q}_{\nu_{k+1}}(k+1)\boldsymbol{A}(k+1) - \boldsymbol{Q}_{\nu_k}(k) &= \boldsymbol{Q}_{\nu_{k+1}}(k+1) - \boldsymbol{Q}_{\nu_k}(k) - \nu_{k+1}\boldsymbol{I}_N \\ &\le -\gamma \boldsymbol{I}_N. \end{aligned} \tag{13}$$

This condition is met if

$$\|\boldsymbol{Q}_{\nu_{k+1}}(k+1) - \boldsymbol{Q}_{\nu_k}(k)\| \le \xi \nu_{k+1} \tag{14}$$

where ξ is a real positive constant, $\xi < 1$, and $\|(\cdot)\|$ is the induced L_2 norm of the matrix $(\cdot)$. Dividing both sides of (14) by ν_{k+1}, (14) becomes

$$\|\boldsymbol{Q}_1(k+1) - \nu_{k+1}^{-1}\boldsymbol{Q}_{\nu_k}(k)\| \le \xi < 1 \tag{15}$$

for all k, where $\boldsymbol{Q}_1(k)$ is given by

$$\boldsymbol{Q}_1(k) = \sum_{i=0}^{+\infty} [\boldsymbol{A}^T(k)]^i \boldsymbol{A}^i(k). \tag{16}$$

The above condition can be used to enable BIBO stable operation of an adaptive IIR filter equipped with a projection technique. Any projection technique that can move the coefficients back to a space that satisfies the condition in (15) can be used for this purpose. However, since most projection techniques have unpredictable computational complexity and since they may result in coefficient stalling, we now derive a closed-form bound for μ_k when the parameter $\nu_k = 1$, under the assumption that the coefficient vary slowly. In this case, the inequality in (15) reduces to

$$\|\boldsymbol{Q}_1(k+1) - \boldsymbol{Q}_1(k)\| \le \xi < 1 \tag{17}$$

for all k. Since ν_k is fixed, we drop the subscript on $\boldsymbol{Q}$ such that $\boldsymbol{Q}(k) = \boldsymbol{Q}_1(k)$ from now on.

To derive the result, we note that

$$\|\boldsymbol{Q}(k+1) - \boldsymbol{Q}(k)\| \le \|\text{vec}[\boldsymbol{Q}(k+1)] - \text{vec}[\boldsymbol{Q}(k)]\|. \tag{18}$$

Combining (17) and (18), we derive a sufficient condition for the exponential stability of the system in (5) to be

$$\|\text{vec}[\boldsymbol{Q}(k+1)] - \text{vec}[\boldsymbol{Q}(k)]\| \le \xi < 1 \tag{19}$$

for all k. In the hypothesis of slowly varying coefficients, the following approximation can be applied:

$$\text{vec}[\boldsymbol{Q}(k+1)] - \text{vec}[\boldsymbol{Q}(k)] \simeq \nabla_{\boldsymbol{\theta}} \text{vec}[\boldsymbol{Q}(k)] \cdot \Delta\boldsymbol{\theta}(k) \tag{20}$$

where $\nabla_{\boldsymbol{\theta}}$ indicates the gradient vector operator with respect to the coefficient vector $\boldsymbol{\theta}$ and $\Delta\boldsymbol{\theta}(k) = \boldsymbol{\theta}(k+1) - \boldsymbol{\theta}(k)$. Recall from (4) that $\Delta\boldsymbol{\theta}(k) = \mu_k \boldsymbol{\psi}(k)$. Substituting the approximation of (20) in (19) and manipulating the resulting expression gives an explicit condition on μ_k for the stability of (5) to be

$$\mu_k \le \frac{\xi}{\|\nabla_{\boldsymbol{\theta}} \text{vec}[\boldsymbol{Q}(k)] \cdot \boldsymbol{\psi}(k)\|} \tag{21}$$

for all k.

The stability conditions of (15) and (17) are derived without resorting to any approximation. However, these conditions can be employed in adaptive recursive filtering applications only with the help of projection techniques. Even though the derivation of (21) employs an approximation that is based on slow variations in the coefficients, this condition has the advantage of being useful in directly controlling the step size of adaptation.

1) Second-Order Case: The stability conditions derived in the previous subsection hold for any filter order. Computationally simple expressions that relate the filter coefficients to the stability bound can be derived for the second-order case. Therefore, the implementation of the stability conditions is simplest when the adaptive filter is realized as a cascade or parallel connection of second-order sections. Since the nonrecursive part does not affect the stability of the resulting system, provided that the coefficients are bounded, we consider the following second-order filter:

$$y(k) = a_1(k)y(k-1) + a_2(k)y(k-2) + x(k). \tag{22}$$

For the instantaneous poles of this system to be inside the unit circle, the coefficients must satisfy the inequalities

$$|a_1(k)| + a_2(k) < 1 \tag{23}$$

and

$$a_2(k) > -1. \tag{24}$$

The candidate Lyapunov matrix $Q(k)$ in (16) can be shown to be given by

$$Q(k) = \begin{bmatrix} 2\,\dfrac{a_2(k)-1}{r(k)} & -2\,\dfrac{a_1(k)a_2(k)}{r(k)} \\ -2\,\dfrac{a_1(k)a_2(k)}{r(k)} & \dfrac{s(k)}{r(k)} \end{bmatrix} \tag{25}$$

where

$$\begin{aligned} r(k) &= -a_2^3(k) + a_2^2(k) + a_1^2(k)a_2(k) \\ &\quad + a_2(k) + a_1^2(k) - 1 \\ &= (a_2+1)[-(a_2-1)^2 + a_1^2] \end{aligned} \tag{26}$$

and

$$\begin{aligned} s(k) &= a_2^3(k) - a_2^2(k) + a_1^2(k)a_2(k) \\ &\quad + a_2(k) + a_1^2(k) - 1. \end{aligned} \tag{27}$$

Substituting (25) into (21) results in the bound for μ_k for the second-order system in (28), shown at the bottom of the page, where

$$\begin{aligned} w(k) &= (3a_2^2(k) + a_1^2(k) - 2a_2(k) + 1)\psi_2(k) \\ &\quad + 2a_1(k)(a_2(k)+1)\psi_1(k) \end{aligned} \tag{29}$$

and

$$\begin{aligned} v(k) &= (-3a_2^2(k) + 2a_2(k) + a_1^2(k) + 1)\psi_2(k) \\ &\quad + 2a_1(k)(a_2(k)+1)\psi_1(k). \end{aligned} \tag{30}$$

$$\mu_k \le \frac{r^2(k)}{\sqrt{4(r(k)\cdot\psi_2(k) - (a_2(k)-1)\cdot v(k))^2 + 8((a_1(k)\psi_2(k) + a_2(k)\psi_1(k))\cdot r(k) - a_1(k)a_2(k)\cdot v(k))^2 + (r(k)\cdot w(k) - v(k)\cdot s(k))^2}} \tag{28}$$

It follows from (26), (24), and (23) that $r(k)$ cannot be zero if the instantaneous poles of the system in (22) are inside the unit circle.

We can see from (28) as well as (21) that the magnitude of the coefficient increment does not depend on $\psi(k)$ as the magnitude of $\psi(k)$ increases to infinity when the step size is controlled by our approach. Consequently, coefficient updates are not stalled when $\psi(k)$ becomes large. The magnitude of the coefficient increment vector depends primarily on the location of the instantaneous poles of the adaptive filter.

III. Stabilized Output Error Adaptive IIR Filters

In this section, we apply the stability condition derived in Section II to the stabilization of output-error adaptive IIR filters. Even though the ideas presented here are applicable to almost all adaptive IIR filters, we describe our approach using the Gauss-Newton output error adaptation algorithm [9]. Furthermore, since the implementation of the stability condition is simplest when the adaptive filter is realized as a cascade or parallel connection of second-order sections, we have considered adaptive IIR filters employing parallel second-order sections.

Each second-order section is of the form

$$\begin{aligned} y_i(k) &= a_{1i}(k)y_i(k-1) + a_{2i}(k)y_i(k-2) \\ &\quad + b_{0i}u(k) + b_{1i}u(k-1) + b_{2i}u(k-2) \end{aligned} \tag{31}$$

where $u(k)$ denotes the input to the section. Let the data vector and the coefficient vector of the ith section be given by

$$\boldsymbol{X}_i(k) = [u(k), u(k-1), u(k-2), y_i(k-1), y_i(k-2)]^T \tag{32}$$

and

$$\boldsymbol{\theta}_i(k) = [b_{0i}(k), b_{1i}(k), b_{2i}(k), a_{1i}(k), a_{2i}(k)]^T \tag{33}$$

respectively. We define the data and coefficient vectors for the overall structure to be

$$\boldsymbol{X}(k) = [\boldsymbol{X}_1^T(k), \boldsymbol{X}_2^T(k), \cdots, \boldsymbol{X}_L^T(k)]^T \tag{34}$$

and

$$\boldsymbol{\theta}(k) = [\boldsymbol{\theta}_1^T(k), \boldsymbol{\theta}_2^T(k), \cdots, \boldsymbol{\theta}_L^T(k)]^T \tag{35}$$

respectively, where L denotes the number of parallel sections.

The coefficients are updated in this method as

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \boldsymbol{\mu}(k)\boldsymbol{R}^{-1}(k+1)\boldsymbol{\phi}(k)e(k) \tag{36}$$

where $\boldsymbol{\mu}(k)$ is a time-varying step size matrix of the adaptive filter defined as

$$\boldsymbol{\mu}(k) = \mathrm{diag}[\mu_1(k), \mu_2(k), \cdots, \mu_{5L}(k)] \tag{37}$$

where $e(k)$ is the *a priori* estimation error defined as

$$e(k) = d(k) - \sum_{i=1}^{L} y_i(k) \tag{38}$$

and

$$\boldsymbol{\phi}(k) = [\boldsymbol{\phi}_1^T(k), \boldsymbol{\phi}_2^T(k), \cdots, \boldsymbol{\phi}_L^T(k)]^T \tag{39}$$

is the information vector whose ith vector element $\boldsymbol{\phi}_i(k)$ is

$$\boldsymbol{\phi}_i(k) = \frac{1}{1 - a_{1i}(k)q^{-1} - a_{2i}(k)q^{-2}}\boldsymbol{X}_i(k). \tag{40}$$

In the above expression, the notation q^{-1} refers to a unit delay operator. The matrix $\boldsymbol{R}(k)$ is an estimate of the autocorrelation matrix of the information vector, and it is recursively computed as

$$\boldsymbol{R}(k) = \lambda\boldsymbol{R}(k-1) + (1-\lambda)\boldsymbol{\phi}(k)\boldsymbol{\phi}^T(k) \tag{41}$$

where $0 \ll \lambda < 1$ is a parameter that controls the convergence and tracking speed of the estimation of the autocorrelation matrix. Its inverse may be evaluated recursively using the matrix inversion lemma as

$$\begin{aligned} &\boldsymbol{R}^{-1}(k+1) \\ &= \frac{1}{\lambda}\left(\boldsymbol{R}^{-1}(k) - \frac{\boldsymbol{R}^{-1}(k)\boldsymbol{\phi}(k)\boldsymbol{\phi}^T(k)\boldsymbol{R}^{-1}(k)}{\frac{\lambda}{1-\lambda} + \boldsymbol{\phi}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{\phi}(k)}\right). \end{aligned} \tag{42}$$

The BIBO stability of the above adaptive filter can be achieved by constraining the step sizes associated with the recursive component of each section to meet the conditions specified by (28). We point out again that the instantaneous poles must always lie within the unit circle and that the coefficients of the feedforward part must be bounded. In all our experiments, we have also limited the maximum step size value. Doing so has two advantages: i) It allows the designer to control the steady-state behavior of the adaptive filter independently of the characteristics of the adaptive filter coefficients, and ii) it ensures that the coefficients vary slowly so that the approximations in the derivation are valid.

The computational complexity of calculating the step size bound in (28) corresponds to 16 multiplications, one square-root operation, and one division per second-order section. Consequently, the complexity of implementing the stability bounds for a cascade or parallel adaptive filter is linearly proportional to the order of the filter. Furthermore, this complexity is comparable with or smaller than the complexity of adapting the coefficients in many adaptive IIR filtering algorithms.

IV. Experimental Results

In the first set of results presented below, the adaptive filters were employed to identify an unknown, fourth-order IIR filter with transfer function

$$\begin{aligned} H(z) = &\frac{1}{1 - 1.86z^{-1} + 0.8698z^{-2}} \\ &+ \frac{2}{1 - z^{-1} + 0.5z^{-2}} \end{aligned} \tag{43}$$

using measurements of the input and output signals. The poles of the above system are located at $[0.93 \pm 0.07j]$ and $[0.5 \pm 0.5j]$. The adaptive filters employed a parallel connection of two second-order systems and were adapted using the Gauss–Newton algorithm. The input of the unknown system is a colored Gaussian signal with zero mean value obtained by filtering a white Gaussian signal with zero mean value and unit variance with the FIR filter of transfer function given by

$$W(z) = 1 + 0.5z^{-1}. \tag{44}$$

The desired response signal was generated by processing this signal with the unknown system and then corrupting the output with an additive, zero-mean and white Gaussian noise sequence that is statistically independent of the input signal. The variance of the measurement noise was such that the output signal-to-noise ratio was 30 dB. The adaptive filter employed a different step size sequence for each second-order section and for the recursive and nonrecursive part of each section. The step size of the recursive part was selected to be the minimum of 0.001 or the bound suggested by our conditions, whereas that of the moving average part was fixed at 0.0005. The forgetting factor in the evaluation of the inverse of the autocorrelation matrix was chosen to be 0.9999. Almost all output error adaptive recursive filters are susceptible to

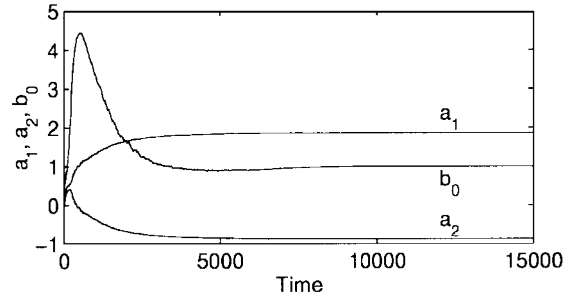

Fig. 1. Evolution of the feedback coefficients of one of the parallel sections of the stabilized adaptive IIR filters.

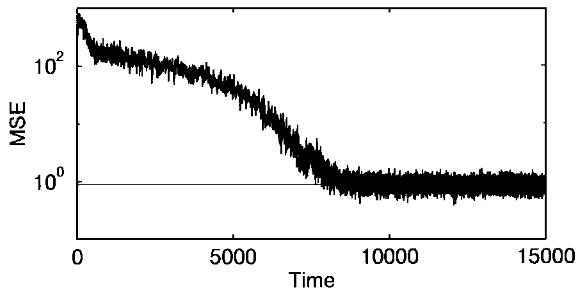

Fig. 2. Evolution of the mean-square estimation error in the stabilized adaptive IIR filters.

converging to the wrong local minima of the squared estimation error surface. In the results presented here, all the experiments that resulted in convergence to wrong local minima were eliminated from the calculation of the ensemble averages. In this way, we are able to observe the speed of convergence of the adaptive filter when it converged to the true solution. The results displayed in the figures are averages of the first 50 experiments in which the coefficients converged to the correct solution.

In addition to constraining the step size to values below the stability bound at each time, we must also verify that the instantaneous poles of the updated filter are within the unit circle. In the experiments described below, the updates for a particular iteration were simply skipped whenever one or more poles crossed the unit circle. Excursions of the poles outside the unit circle have occurred rarely in our experiments when the step size was selected according to our bounds. Consequently, problems due to coefficient stalling because of skipped updates did not occur in any of our experiments. In order to ensure the BIBO stability of the adaptive filter feedforward part, we also imposed an upper bound on the absolute value of feedforward coefficients. The upper bound chosen was 1000, and in no experiment did the feedforward coefficients reach this bound. The algorithm was initialized with the coefficients of the feedforward parts equal to 0.5 and the poles of the recursive part equal to [$0.1 \pm 0.1i$] and [-0.1 ± 0.1i], respectively, for each second-order section.

Fig. 1 shows the ensemble averaged behavior of the coefficients of the parallel section that correspond to the coefficients 1.86 and 0.8698 (corresponding to the poles located at [$0.93 \pm 0.07j$]) of the unknown system. Fig. 2 shows the ensemble averaged, squared estimation error at the output of the adaptive filters. The horizontal line in the figure represents the noise floor. Fig. 3 displays the ensemble averaged step size sequence for the parallel section tracking the poles of the unknown system at [$0.93 \pm 0.07j$]. The results indicate that step size selection using the closed-form conditions in (28) results in stable operation of the adaptive filter. The initial values of the step size are small in this example because the initial estimation error is large. Combined with the large error, the initial values of the step size produced the largest changes possible that still maintained the exponential stability of the system.

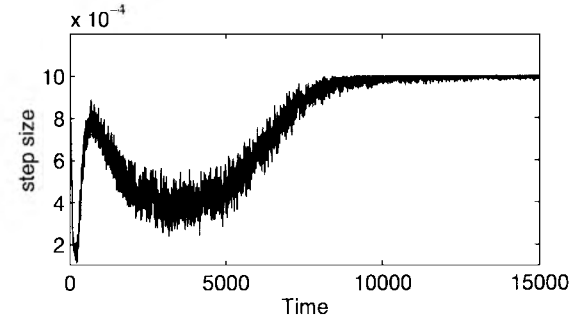

Fig. 3. Evolution of the step size sequence of the stabilized adaptive IIR filters.

We now compare the performance of the stabilized adaptive IIR filter with that of the SHARF algorithm. In order to make the comparisons as fair as possible, we used a single second-order system for identifying an unknown second-order system with transfer function

$$H(z) = \frac{1}{1 - 1.9z^{-1} + 0.905z^{-2}}. \tag{45}$$

We used the same experimental conditions as in the previous example, with the difference that we employed the same step size sequence for adapting the moving average and the recursive coefficients of the system. The coefficient update in the SHARF algorithm was implemented as in [15] in the following manner:

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \mu \boldsymbol{R}^{-1}(k+1)\boldsymbol{X}(k)e_f(k) \tag{46}$$

where

$$\boldsymbol{R}^{-1}(k+1) = \frac{1}{\lambda}\left(\boldsymbol{R}^{-1}(k) - \frac{\boldsymbol{R}^{-1}(k)\boldsymbol{X}(k)\boldsymbol{X}^T(k)\boldsymbol{R}^{-1}(k)}{\dfrac{\lambda}{1-\lambda} + \boldsymbol{X}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{X}(k)}\right) \tag{47}$$

$$e(k) = d(k) - \boldsymbol{\theta}^T(k)\boldsymbol{X}(k) \tag{48}$$

and

$$e_f(k) = \sum_{m=0}^{p-1} c(m)e(k-m). \tag{49}$$

The input vector $\boldsymbol{X}(k)$ is defined in this case as $[x(k), y(k-1), y(k-2)]^T$. In order to obtain satisfactory convergence of the SHARF algorithm, the FIR filter with transfer function $C(z)$ that gives the filtered estimation error $e_f(k)$ must be such that $C(z) \simeq A(z)$, where $A(z)$ is the denominator of the transfer function of the unknown filter. In our case, we have considered the optimal choice of $C(z) = A(z)$.

Both the algorithms were initialized with the coefficient of the feedforward part equal to 0.5 and with the two poles at the origin.

The step size μ was selected to be 0.000 08 for the SHARF algorithm so that the steady-state excess mean-square error was identical to that of the stabilized adaptive IIR filter of this paper. We note that the coefficient update equations (46) and (47) have the form of a Gauss–Newton update. The similarity of this set of update equations to those in (36)–(42) and the choices of the step sizes so that the steady-state errors are almost identical make it possible to make fair comparisons of the performance of the two algorithms.

Fig. 4 plots the evolution of the mean-squared estimation error for the two algorithms. We can see from this figure that the SHARF

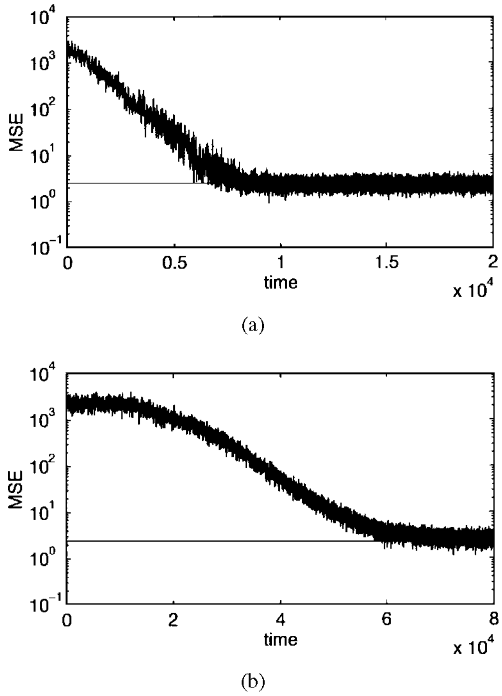

Fig. 4. Comparison of the stabilized adaptive filter of this paper and the SHARF algorithm.

algorithm converges much slower than the method introduced in this correspondence. Note that the time scales used in the two plots are different from each other. In general, when the instantaneous poles are initialized to be sufficiently removed from the unit circle, we have observed that our method converges significantly faster than the SHARF. However, it is possible to slow the initial convergence rate of our method by initializing the instantaneous poles to be very close to the unit circle. Such an initialization will force the initial values of the step size to be very small and, therefore, will result in slow convergence.

We also studied the convergence behavior of the Gauss–Newton output error algorithm with fixed step size using the same experimental conditions. Without pole projection inside the unit circle, the algorithm became unstable in all 50 experiments we performed with a step size equal to $6 \cdot 10^{-4}$, which was lower than the maximum step size value we allowed for the stabilized adaptive filter. With pole projection and a step size equal to $8 \cdot 10^{-4}$, the instantaneous poles moved to locations outside the unit circle so often that the overall speed of convergence was much slower than that of the stabilized algorithm. Furthermore, the evolution of the coefficients toward their steady-state values was very erratic for the method employing only pole projection. In our experiments, with the adaptive filter employing the new step-size bound, only in one of the 51 experiments did the coefficients of the system not converge to the correct coefficient values after 20 000 samples. With the fixed step size, the coefficients in 16 of 66 experiments did not converge to the correct values during the same time span.

Finally, we compare the bounds given by (28) with the bounds derived in [1] and [2] for the maximum allowable variations in the coefficients of an exponentially stable, second-order linear system with time-varying coefficients. The stability bounds in [1] and [2] are expressed in terms of the state transition matrix. In order to make the comparison as fair as possible, we derived the maximum allowable coefficient variation for the state transition matrix defined in (9). Fig. 5 displays the three bounds as a function of the magnitude of the complex instantaneous poles of the system. The curve a refers

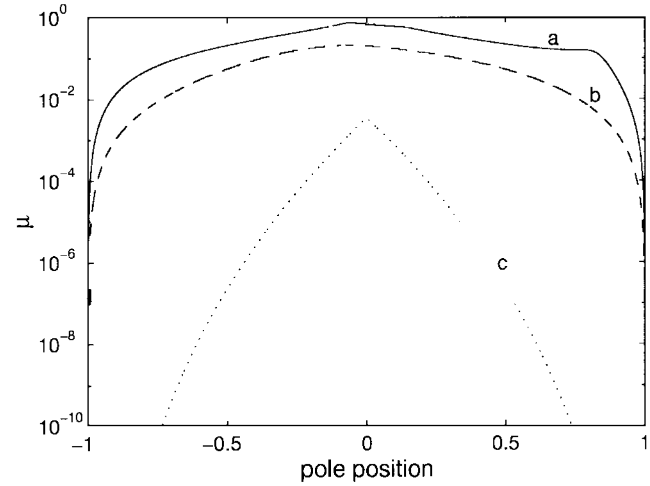

Fig. 5. Comparison of the bound in (28) with that derived in [1] and [2].

to the bound given by (28), whereas the curves b and c refer to the stability bounds derived in [1] and [2], respectively.

The bounds were obtained for the case when the complex conjugate pole pair moved along straight lines located at $\pm 45°$ to the real axis. The coefficients were assumed to change as in (4), and the vector $\phi(k)$ was assumed to have unit magnitude at each time. The bounds are for the scaling factor μ_k in the coefficient evolution equation (4). It is clear from the results of Fig. 5 that all three stability bounds converge to zero as the instantaneous poles approach the unit circle. However, the rate at which $\mu(k)$ in (28) approaches zero when the poles tend to the unit circle is several orders of magnitude slower than the bound derived in [2]. Furthermore, the bound for $\mu(k)$ given by (28) is much greater than the bound in [1]. This result is an additional demonstration of the usefulness of the sufficient stability bounds derived in this correspondence.

V. Concluding Remarks

This correspondence presented a novel stability condition for time-varying direct-form recursive linear systems. This condition was successfully applied for designing bounded-input, bounded-output stable adaptive IIR filters. The experimental results not only confirmed the reliability of the derived bound but also demonstrated the better convergence characteristics of the stabilized algorithm when compared with other stable adaptive IIR filters. The time-varying bound derived in this correspondence may be incorporated into any practical adaptive IIR filter. It is well known that certain adaptive IIR filtering algorithms such as Feintuch's method diverge for all choices of the step size for certain input signals [16]. Experimental results, as well as theoretical considerations, indicate that the step-size bound derived in this correspondence eventually goes to zero in such situations, thus preserving the BIBO stability of the adaptive filter.

References

[1] F. Amato, G. Celentano, and F. Garofalo, "New sufficient conditions for the stability of slowly-varying linear systems," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 1409–1411, Sept. 1993.

[2] C. A. Desöer, "Slowly-varying discrete system $x_{i+1} = A_i x_i$," *Electron. Lett.*, vol. 6, no. 11, pp. 339–340, May 1970.

[3] P. L. Feintuch, "An adaptive recursive LMS filter," *Proc. IEEE*, vol. 64, pp. 1622–1624, Nov. 1976.

[4] C. R. Johnson, Jr., "A convergence proof for a hyperstable adaptive recursive filter," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 745–749, Dec. 1979.

[5] ——, "Adaptive IIR filtering: Current results and open issues," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 237–250, Mar. 1984.

[6] G. Kubin, "On the stability of wave digital filters with time-varying coefficients," in *Proc. Euro. Conf. Circuit Theory Design*, Prague, Czechoslovakia, Sept. 2–6, 1985, pp. 499–502.
[7] I. D. Landau,"Unbiased recursive identification using model reference adaptive techniques," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 194–202, Apr. 1976.
[8] M. G. Larimore, J. R. Treichler, and C. R. Johnson, Jr., "SHARF: An algorithm for adapting IIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 428–440, Aug. 1980.
[9] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge, MA: MIT Press, 1983.
[10] K. X. Miao, H. Fan, and M. Doroslovački, "Cascade normalized lattice adaptive IIR filters," *IEEE Trans. Signal Processing*, vol. 42, pp. 721–742, Apr. 1994.
[11] S. L. Netto, P. S. R. Diniz, and P. Agathoklis, "Adaptive IIR filtering algorithms for system identification: A general framework," *IEEE Trans. Educ.*, vol. 38, pp. 54–66, Feb. 1995.
[12] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. New York: Marcel Dekker, 1995.
[13] H. H. Rosenbrock, "The stability of linear time-dependent control systems," *J. Electron. Contr.*, vol. 15, pp. 73–80, July 1963.
[14] W. J. Rugh, *Linear System Theory*. Upper Saddle River, NJ: Prentice-Hall, 1995.
[15] J. J. Shynk, "Adaptive IIR filtering," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 6, pp. 4–21, Apr. 1989.
[16] G. A. Williamson, B. D. O. Anderson, and C. R. Johnson, Jr., "On the local stability properties of adaptive parameter estimators with composite errors and split algorithms," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 463–473, Apr. 1991.

A New Method for Varying Adaptive Bandwidth Selection

Vladimir Katkovnik

***Abstract*—A novel approach is developed to solve a problem of varying bandwidth selection for filtering a signal given with an additive noise. The approach is based on the intersection of confidence intervals (ICI) rule and gives the algorithm, which is simple to implement and adaptive to unknown smoothness of the signal.**

***Index Terms*—Adaptive filtering, adaptive varying bandwidth, adaptive varying window length, segmentation.**

I. Introduction

In this correspondence, we introduce an adaptive filter that produces piecewise smooth curves with a small number of discontinuities in the signal or its derivatives. It allows certain desirable features such as jumps or instantaneous slope changes to be preserved in the smooth curves. The algorithm is adaptive to unknown smoothness of the signal. The local polynomial approximation (LPA) is used as a tool for filter design as well as for a presentation of the developed general method of the bandwidth selection. This method can be applied for a variety of quite different linear and nonlinear problems where the bandwidth selection involves the bias-variance compromise usual for nonparametric estimation.

Manuscript received January 20, 1998; revised January 26, 1999. This paper was supported in part by the Foundation of Research Development of South Africa. The associate editor coordinating the review of this paper and approving it for publication was Dr. Phillip A. Regalia.

The author is with the Department of Statistics, University of South Africa, Pretoria, South Africa (e-mail: vlkatkov@hotmail.com).

Publisher Item Identifier S 1053-587X(99)06767-7.

TABLE I
Square Root Mean Squared Errors (SRMSE) of Estimation Using the LPA and Various Wavelet Methods

N	$\Gamma = var$	$\Gamma = 4.4$	$Wavelets$
Blocks			
256	0.61	1.56	(0.68-1.20)
512	0.46	0.92	(0.59-1.12)
1024	0.33	0.74	(0.47-1.03)
2048	0.25	0.49	(0.41-0.85)
Heavy			
256	0.47	0.79	(0.49-0.62)
512	0.37	0.67	(0.40-0.55)
1024	0.31	0.49	(0.32-0.46)
2048	0.24	0.34	(0.38-0.61)

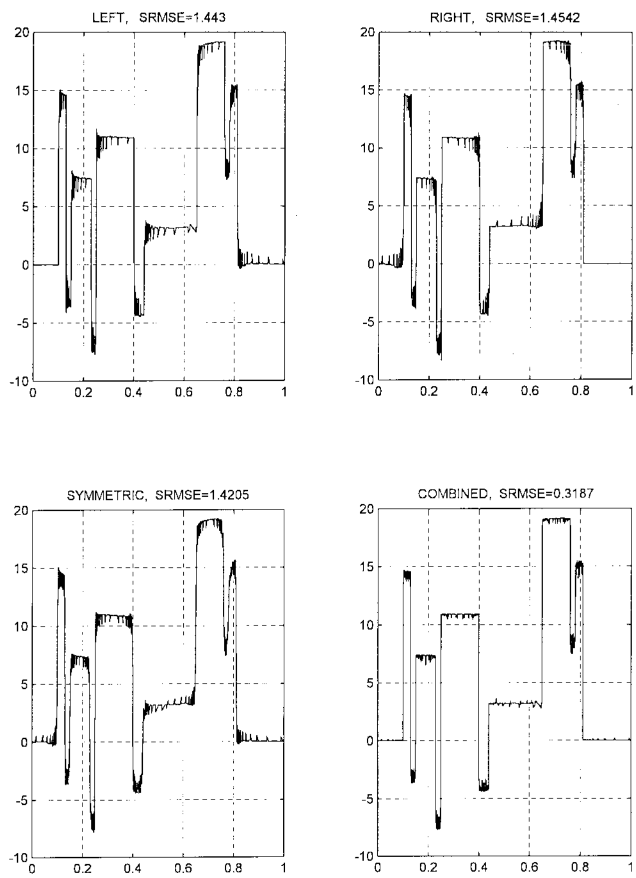

Fig. 1. Adaptive estimates, $m = 2$, $\Gamma = 4.4$.

Suppose that we are given noisy observations of a signal $y(x)$ with a sampling period Δ, i.e., $x_s = s\Delta$, $s = \cdots - 1, 0, 1 \cdots$: $z_s = y(x_s) + \varepsilon_s$, where ε_s are independent and identically distributed Gaussian random errors $E(\varepsilon_s) = 0$, $E(\varepsilon_s^2) = \sigma^2$. It is assumed that